

GINTIS, Herbert, 2009, *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*, Princeton, New Jersey, Princeton University Press. 304 pp.

Herbert Gintis es un economista estadounidense cuyos trabajos abarcan un amplio espectro de temas relacionados a varias de las llamadas ciencias del comportamiento (antropología, biología, ciencia política, economía, sociología, sicología). En particular, *The Bounds of Reason* representa un intento por unificar estas ciencias bajo un marco conceptual común provisto por la teoría de juegos. El principal argumento que desarrolla Gintis en este libro es que la teoría de juegos necesita de una teoría social más amplia para lograr mayor poder explicativo, al mismo tiempo que la teoría social sin la teoría de juegos resulta incompleta para explicar el comportamiento humano. Para persuadirnos de ello, Gintis discute con rigor los principales conceptos usados en la teoría de juegos; muestra, cuando es pertinente, sus limitaciones para predecir el comportamiento humano; y, luego de demostrar la incompatibilidad entre los modelos de comportamiento individual asumidos en la biología, la economía, la psicología y la sociología, propone herramientas analíticas que podrían ayudar a unificar dichas ciencias en un solo modelo de elección e interacción estratégica.

The Bounds of Reason es un texto accesible para alumnos avanzados de pregrado o de primeros ciclos de postgrado que dominan la teoría de juegos y alguna noción de economía experimental (Friedman y Sunder) y de economía del comportamiento (Gilovich *et al.*), por lo cual en esta reseña se asume cierta familiaridad con dichas áreas. El libro está dividido en trece capítulos; los dos primeros introducen nociones de microeconomía neoclásica (preferencias consistentes, utilidad esperada), economía del comportamiento (teoría prospectiva y algunos sesgos en las decisiones realizadas en un contexto de incertidumbre o riesgo: representatividad, disponibilidad, anclaje e imitación) y teoría de juegos (formas normal y extensa, estrategias mixtas, equilibrio de Nash y equilibrio correlacionado) y aplicaciones de la teoría de juegos, presentadas a un nivel similar al de Gibbons.

En el capítulo 3, Gintis examina el aporte de la teoría de juegos en el diseño de experimentos conducidos para entender el comportamiento humano en situaciones de interacciones estratégicas. Aquí se empieza a discutir la evidencia acerca de las predicciones de la teoría de juegos cuando se asume que las preferencias de los agentes económicos son egoístas (*self-regarding*). Sugestivamente, Gintis reporta que dichas predicciones son correctas cuando se modelan mercados con contratos bien especificados (como en las subastas dobles y en los oligopolios) donde el intercambio es anónimo; no obstante, cuando los contratos son incompletos, y los agentes pueden interactuar

estratégicamente y tienen la potestad de premiar y castigar el comportamiento de otros agentes, las predicciones de la teoría de juegos que mantienen ese supuesto generalmente fallan. Este es un resultado significativo encontrado por varios estudios experimentales, en los que se verifica que preferencias que se preocupan por el prójimo (*other-regarding preferences*), como el altruismo, la cooperación, la aversión a la desigualdad y la reciprocidad, explican el comportamiento de los agentes bajo distintos contextos de interacciones individuales y sociales.

Los capítulos 4 a 10 son más intensivos en su tratamiento de tópicos de la teoría de juegos, presentados a un nivel comparable al de Osborne y Rubinstein. Los capítulos 4 y 5 discuten los temas de la *rationalizability* (que no parece tener traducción al castellano) y la racionalidad del conocimiento público (*common knowledge rationality*, CKR) aplicadas a la teoría de juegos. En ese contexto, un agente racional no escogerá un curso de acción o una estrategia que es dominada (es decir, que le reporta menores ganancias que otras estrategias, dados los planes de acción de los demás agentes). Asimismo, la CKR implica que todos los agentes saben que son racionales y que todos ellos saben que todos saben que lo son. Gintis cuestiona la validez de este supuesto e indica que la CKR es, más bien, un evento en el sentido estadístico del término: se cumple en algunas circunstancias.

El capítulo 6 discute la inexistencia de incentivos compatibles que se da cuando, en equilibrio, los agentes escogen estrategias mixtas (*i. e.*, ninguna estrategia es escogida con probabilidad igual a 1) y critica la racionalidad detrás de la existencia de equilibrios en estrategias mixtas. Este capítulo aparece algo desconectado del resto. Los dos capítulos siguientes introducen la importancia de las creencias en el análisis del equilibrio en la teoría de juegos y nos van presentando nociones como las normas sociales, que pueden surgir en equilibrio como instancias de coordinación de las estrategias de los agentes (esto vendría a ser un equilibrio correlacionado, donde existe un «coreógrafo» que coordina las acciones, citando a Gintis). El capítulo 8 llama la atención de lo beneficioso de relacionar conceptos usados en la economía –como el del comportamiento racional– con el de instituciones sociales, lo que puede ser considerado para alinear las creencias y expectativas de los individuos y regular desde las relaciones familiares hasta la división del trabajo.

En el capítulo 9, Gintis plantea un refinamiento del equilibrio para juegos en forma extensa, llamado criterio de mejor respuesta local, el cual, según él argumenta, es superior a los planteados en la literatura. No queremos entrar en detalles al respecto, baste con mencionar que Gintis muestra ejemplos en los cuales este criterio predice resultados mejor que los refinamientos más comunes (*e. g.*, equilibrio perfecto, bayesiano perfecto y secuencial).

Los tres capítulos siguientes representan el corazón del análisis de los aspectos relacionados a las interacciones sociales y de posibles sinergias entre las antes mencionadas ciencias del comportamiento. En particular, el capítulo 10 discute las condiciones para la existencia de cooperación bajo distintos escenarios en el contexto de la teoría de juegos. En el capítulo 11, de manera interesante, Gintis argumenta que el efecto de certidumbre (*endowment effect*) propuesto por la teoría prospectiva es la base de varias formas de socialidad, incluyendo el respeto de los derechos de propiedad. De esta manera, los propietarios están dispuestos a gastar más recursos con tal de proteger su propiedad que un intruso con tal de apropiarse de ella. El capítulo 12 concluye que un modelo de teoría de juegos que explique la cooperación debiera contar con un modelo psicológico de preferencias sociales que incluya el rol de las normas sociales como mecanismo, el cual debe hacer posible que las estrategias de los agentes estén correlacionadas. Además, Gintis propone cinco herramientas analíticas que, argumenta, pueden servir de base común para las ciencias del comportamiento, a saber: la coevolución genético-cultural, la teoría sociopsicológica de las normas, la teoría de juegos, el modelo del actor racional y la teoría de la complejidad. Dependiendo de la disciplina de formación de cada uno, al menos dos de esos términos deben sonar extraños para muchos de nosotros, lo cual probablemente responde a nuestra ignorancia de las interacciones posibles que nuestras disciplinas de formación pueden tener con ramas afines.

En suma, considero que un aporte fundamental de *The Bounds of Reason* es llamar la atención de la importancia, y los beneficios, que un mayor intercambio interdisciplinario podría representar para las ciencias del comportamiento, entre las cuales se encuentra la economía. A diferencia de otros autores, Gintis no cuestiona la validez del modelo neoclásico de agente racional (y egoísta), sino que plantea que el concepto de racionalidad también debería incluir las preferencias que se preocupan por el prójimo y, más importante aun, sugiere que la existencia de anomalías, como la aversión a la pérdida o las preferencias temporales hiperbólicas, no significa que los agentes son irracionales (en el sentido de no tener preferencias consistentes), sino que ellas pueden ser acomodadas dentro de la teoría neoclásica convencional con la ayuda de elementos de las otras ciencias del comportamiento, de tal manera que tales aportes puedan contribuir a que se explique mejor el comportamiento humano.

Francisco B. Galarza
Universidad del Pacífico